Coastal Engineering Technical Note

RELIABILITY OF LONG-TERM WAVE CONDITIONS PREDICTED WITH DATA SETS OF SHORT DURATION

PURPOSE: To provide guidelines for determining the validity and reliability of predicted probable wave heights obtained from data of limited duration.

BACKGROUND: The basic steps listed by Issacson and Mackenzie (1981) provide a good outline for the extrapolation of long-term wave conditions. These steps are listed below.

1. Data made up of significant wave heights and periods are collected or hindcast over a period of time at the site of interest.

2. A plotting formula is used to estimate the long-term probability distribution of wave heights.

3. The results from step 2 are plotted on probability scales corresponding to several given probability distributions. Some of the most commonly used distributions are the Extremal Types I and II, the Weibull, and the Lognormal.

4. For the probability scale that exhibits the best linear plot, a least squares line is fitted through the points to represent the probability distribution for long-term wave conditions.

5. Values of significant wave height or related measures are read from the plotted line for return periods beyond the extent of the observed data. These values are called extrapolated or estimated long-term wave conditions.

DISCUSSION: An important part of selecting the data with which to perform a prediction of long-term wave conditions is the definition of a specific population of interest. In most cases, the wave conditions to be predicted are associated with the more severe climatological events that occur in the study area. The climatological events may be from one or more storm populations

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **MAR 1985**	2. REPORT TYPE	3. DATES COVERED **00-00-1985 to 00-00-1985**
4. TITLE AND SUBTITLE **Realibility Of Long-Term Wave Conditions Predicted With Data Sets Of Short Duration**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Engineer Waterways Experiment Station,Coastal Engineering Research Center,Vicksburg,MS,39180**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

such as hurricanes, tropical storms, and winter storms. Wave conditions resulting from different storm populations must be analyzed separately to avoid the effects of mixing on the results. To predict long-term extreme wave conditions in an area where hurricanes are the meteorological event of major severity, hurricane data must be used to make the prediction. If wave data resulting from hurricanes are mixed with wave data resulting from other meteorological conditions, then invalid predictions can occur. Another example of this type error is in areas where there is a winter storm season. If monthly maxima from a 2-year record of wave height data were used to make a long-term prediction, then the results would be biased by the maxima from months during which no storms occurred; thus, only data from the storm season are applicable to the extremal analysis. A good discussion of the problems with mixing populations is given by Borgman and Resio (1977).

Once the very important step of defining the storm type populations has been completed, then the available data set can be surveyed for wave conditions that correspond to desired storm events. Measured data sets that include wave conditions from 1 to 5 years usually do not include enough storm events to perform a reliable prediction of long-term wave conditions. The general rule of thumb is that it is not good to extrapolate to more than three times the extent of the data set (Borgman and Resio 1977). Following this rule of thumb, if a 1-year data set provided enough events of interest to perform the analysis outlined in steps 2 to 5, the prediction would only be reliable for up to a 3-year return period. For a 5-year data set, the longest return period condition that could be reliably predicted would be the 15-year condition. It is evident that to reliably predict the usual 50-year wave condition used for design purposes would require at least 17 years of data. The Wave Information Study (WIS) at CERC is a good source for long-term hindcast data. The data retrieval and analysis program known as the Sea State Engineering Analysis System (SEAS) makes handling of the available hindcast data convenient. The SEAS report (WIS Report 10, Ragsdale 1983) provides a listing of all hindcast locations and data retrieval and analysis procedures available to the user.

Wave data records covering 17 or more years are usually only available as hindcast data, derived from synoptic weather records. For this reason, hindcast data are the most useful for making long-term predictions.

For situations in which hindcast data are not available and long-term predictions must be computed from the more limited measured wave records, the desired statistical reliability cannot be obtained. The general reliability of long-term predictions computed with measured data can be increased by using collateral information such as long-term weather maps, visual observations from long-term local residents, and long-term wind records. In some cases, the extremes may be limited by water depth or geographical factors that can be used to simplify the prediction. The major point is that extrapolations obtained directly from a measured record of limited extent do not include the important long-term events for the area and must be supplemented by all available long-term information.

Upper 95 percent confidence limits for extrapolated predictions of the standardized Extremal Type I distribution obtained by simulation are listed in Table I. To interpret the numbers in Table I, consider the Extremal Type I cumulative probability function.

$$p = \Pr\left[X \leq x\right] = \exp\left(-\exp\left(\frac{x-u}{\sigma}\right)\right) \tag{1}$$

The numbers in the table correspond to the term

$$z = \left(\frac{x-u}{\sigma}\right) \tag{2}$$

which is the standardized version of the Extremal Type I variable. The steps for using Table I are as follows:

1. The wave height data, x_1, x_2, ..., x_n, are ranked in ascending order.
2. The plotting formula is used to estimate p from equation 1 for each data set.

$$\hat{p} = \frac{r}{n+1}$$

   where r is the rank and n is the sample size.

3. The least squares line $-\ln(-\ln \hat{p}) = Bx + A$ is computed.

Table I

Upper 95 Percent Confidence Limits for the Standardized Extremal Type I Distribution

	CUMULATIVE PROBABILITY (p)														
n	0.9000	0.9200	0.9400	0.9600	0.9800	0.9820	0.9840	0.9860	0.9880	0.9900	0.9920	0.9940	0.9960	0.9980	0.9990
5	5.287	5.744	6.421	7.373	8.983	9.214	9.485	9.799	10.170	10.610	11.133	11.813	12.772	14.387	16.001
10	4.350	4.789	5.307	6.078	7.406	7.607	7.832	8.084	8.368	8.715	9.140	9.649	10.387	11.657	12.964
15	3.920	4.308	4.800	5.487	6.664	6.844	7.043	7.261	7.511	7.807	8.174	8.650	9.323	10.492	11.637
20	3.586	3.959	4.434	5.096	6.164	6.330	6.515	6.727	6.976	7.260	7.615	8.077	8.700	9.788	10.894
25	3.486	3.829	4.248	4.855	5.866	6.025	6.196	6.387	6.619	6.898	7.232	7.647	8.262	9.308	10.336
30	3.298	3.617	4.034	4.626	5.594	5.746	5.919	6.115	6.323	6.575	6.899	7.317	7.882	8.867	9.831
35	3.182	3.488	3.884	4.448	5.399	5.542	5.701	5.882	6.091	6.339	6.642	7.033	7.580	8.524	9.471
40	3.177	3.491	3.888	4.447	5.412	5.558	5.721	5.906	6.119	6.370	6.677	7.074	7.632	8.567	9.524
45	3.090	3.397	3.800	4.353	5.299	5.441	5.600	5.785	6.001	6.252	6.558	6.945	7.486	8.424	9.345
50	3.115	3.424	3.814	4.369	5.305	5.446	5.603	5.783	5.996	6.248	6.556	6.951	7.492	8.417	9.360
55	2.996	3.287	3.677	4.215	5.141	5.276	5.423	5.591	5.794	6.037	6.323	6.691	7.228	8.146	9.063
60	2.957	3.256	3.635	4.167	5.059	5.195	5.343	5.514	5.713	5.948	6.239	6.607	7.129	8.020	8.923
65	2.912	3.207	3.583	4.114	4.994	5.128	5.278	5.451	5.649	5.884	6.171	6.541	7.043	7.940	8.833
70	2.953	3.239	3.620	4.139	5.014	5.147	5.300	5.472	5.671	5.906	6.191	6.553	7.065	7.926	8.796
75	2.906	3.203	3.587	4.095	4.986	5.116	5.258	5.418	5.604	5.830	6.122	6.498	7.028	7.913	8.774
80	2.893	3.172	3.541	4.066	4.949	5.080	5.226	5.396	5.591	5.821	6.101	6.457	6.966	7.831	8.694
85	2.852	3.142	3.500	4.015	4.889	5.021	5.168	5.335	5.524	5.748	6.029	6.391	6.896	7.731	8.602
90	2.807	3.090	3.451	3.970	4.829	4.958	5.103	5.266	5.455	5.680	5.960	6.312	6.808	7.653	8.498
95	2.821	3.105	3.463	3.967	4.826	4.955	5.099	5.262	5.450	5.673	5.944	6.295	6.792	7.643	8.491
100	2.812	3.106	3.454	3.952	4.823	4.952	5.097	5.259	5.445	5.666	5.941	6.295	6.793	7.641	8.494

4. $u = \frac{-A}{B}$, $\sigma = +\frac{1}{B}$

5. To compute the R year predicted significant wave height

find $p = 1 - \frac{1}{R\lambda}$

where R = Return Period

λ = Average number of storms per year (1.0 if using annual maxima)

Then the predicted height is

$$H_{s_R} = \frac{-\ln(-\ln p) - A}{B}$$

6. Choose the table value Z_r with cumulative probability equal to p (use linear interpolation if the required p falls between two table values) and sample size n that is nearest but less than or equal to the data sample size. The upper 95 percent confidence limit for H_{s_r} is

$$CL_R = Z_R \sigma + u$$

Example: Given values for n, A, B, and λ , find the 50-year return period predicted significant wave height and the corresponding upper 95 percent confidence limit (all units are in meters).

Let n = 62 storms (20 years)

λ = 3.1 storms/year

A = -7.567

B = 1.036

Then

$$u = -\left(\frac{-7.567}{1.036}\right)$$

$$= 7.304$$

$$\sigma = \frac{1}{1.036}$$

$$= 0.965$$

$$p = 1 - \frac{1}{50(3.1)}$$

$$= 0.994$$

$$H_{s_{50}} = \frac{- \ell n(- \ell n(0.994) - (- 7.567)}{1.036}$$

$$= 12.24 \text{ m}$$

Look in Table I under $n = 60$ and cumulative probability $= 0.994$ and find $Z_{50} = 6.607$. Then

$$CL_{50} = (6.607)\ (0.965) + 7.304$$

$$= 13.67 \text{ m}$$

If $n = 20$ corresponding to an average of 6.5 years of data, then

$$CL_{50} = (8.077)\ (0.965) + 7.304$$

$$= 15.10 \text{ m}$$

If $n = 10$ for an average of 3.2 years of data, then

$$CL_{50} = (9.649)\ (0.965) + 7.304$$

$$= 16.62 \text{ m}$$

If $n = 5$ for an average of 1.6 years of data, then

$$CL_{50} = (11.813)\ (0.965) + 7.304$$

$$= 18.70 \text{ m}$$

As is evident when the sample size n gets smaller, the upper confidence limits become very large in comparison to the predicted value. This is due to the fact that smaller samples are less likely to contain the larger or more rare extreme events. This causes a higher probability of underestimation to occur; thus, $H_{s_{50}} = 12.24$ m from a sample of 20 years is only 1.5 m less than the 95 percent upper limit. The same prediction with 3.2 years and 1.6 years is 4.4 m and 6.5 m, respectively, less than the 95 percent upper limit.

SUMMARY: An outline of the basic method of extrapolation of long-term wave conditions is given. The reference by Issacson and Mackenzie (1981) is suggested for an in-depth discussion of the methodology.

The validity of the long-term prediction depends heavily upon the population from which the data are taken.

The extent of the data in time determines the reliability of extrapolated predictions. It is an accepted rule of thumb that predictions should not be made beyond three times the extent of the data. Any extrapolation beyond this point may be too unreliable to be useful without some collateral information to support the extrapolation. A table of simulated confidence limits is given along with examples so that the statistical reliability may be approximated and the decision to seek more information or not can be made.

REFERENCES:

Borgman, L. E. and Resio, D. T. 1977. "Extremal Prediction in Wave Climatology," Proceedings, Ports 77, Vol I, New York, N.Y., p 394.

Issacson, M. S. and Mackenzie, N. G. 1981 (May). "Long-Term Distributions of Ocean Waves: A Review," Journal of Waterway, Port, Coastal and Ocean Division, ASCE, Vol 107, No. WW2.

Ragsdale, D. S. 1983 (Aug). "Sea State Engineering Analysis System (SEAS)," Wave Information Studies of US Coastlines (WIS) Report 10, US Army Engineer Waterways Experiment Station, Vicksburg, Miss.